



REMARKS

The present application and its claims are directed to a document ranking system based on user behavior and a document searching system based on user behavior.

RECEIVED  
FEB 24 2003  
Technology Center 2100

PRIOR ART REJECTIONS

In response to the Examiner's rejection of Claims 1- 6 under 35 U.S.C. 102(b) as being anticipated by US Patent No. 5,987,457 to Ballard (hereinafter "Ballard"), and the Examiner's rejection of Claims 7 - 12 under 35 U.S.C. 103(a) as being obvious over Ballard in view of U.S. Patent No. 6,021,409 to Burrows (hereinafter "Burrows"), Applicant respectfully traverses the rejection. In particular, the claims of the application are not anticipated or rendered obvious by the prior art cited by the Examiner for the reasons set forth below. Therefore, Claims 1- 26 are allowable over the prior art cited by the Examiner.

Prior Art

Ballard

Ballard describes a query refinement method for searching documents wherein a user performs a search (using an initial query), views the search results and subjectively determines if a document is desirable or undesirable. (See Abstract.) Using these classifications, the Ballard system extracts a list of keywords from the desirable documents as well as list of keywords from the undesirable documents. The desirable keywords may be weighted so that only the highest weighted keywords may be utilized. The desirable keywords are added to the initial query to form a refined query. The refined query may then be executed so that the search results are more focused due to the addition of the keywords to the initial query. See Col. 2, lines 11 - 24. Thus, in Ballard, the user must explicitly classify documents as either desirable or undesirable and then the system will produce a refined search of the user.

In more detail, Ballard uses the initial query and the keywords ANDed together in order to narrow the search results. See Col. 5, lines 62 - 65. The keywords are generated when the user classifies the documents as desirable or undesirable using a user interface 42. See Col. 6, lines 35 - 50. Ballard requires a user to classify at least one document as "of interest" and at least one document as "not of interest." See Col. 6, lines 40 - 44. Ballard either does not consider uncategorized documents or considers viewed documents as "not of interest." See Col. 6, lines 45 - 50. In other embodiments, Ballard may create a first, second and third set of keywords wherein the

first set of keywords (good keywords) are keywords appearing in the “of interest” document, the second set of keywords (bad keywords) are keywords appearing in the “not of interest” documents and the third set of keywords (dirty keywords) are keywords that appear in both types of documents. See Col. 7, line 26 – Col. 8, line 25.

In Ballard, the system generates a refined query which is formed from the initial query and the good keywords as described above. Thus, Ballard generates refined queries which help the particular user that classified the documents to perform a more focused query. Ballard does not modify/update a feature vector of a document (which would be beneficial to any subsequent user who might perform a search which retrieves the same document).

In summary, Ballard requires explicit user action to refine a query since the user must explicitly classify documents (as “of interest” or “not of interest”) in order to generate the set(s) of keywords that may be used to refine a query. Ballard also requires the generation of a keyword list from the classified documents. Ballard also appears to be a single user system in that the refinement process is applicable to only the user who classifies the documents at the time he does so (rather than for future/subsequent users), since the keywords would only be helpful to the user who classifies the documents and Ballard modifies the query of the user.

#### Burrows

Burrows describes a method for parsing, indexing and searching world-wide-web pages. (See Title) In Burrows, as pointed out by the Examiner, the system uses a “modified collection frequency weighing technique (See Col. 26, lines 28 – 32) in order to rank pages so that pages with a higher rank are presented first to a user. This describes a typical search engine in which each page returned to the user has some ranking and the pages are returned to the user in ranked order.

In Burrows, it appears that a scheme was initially described in which different terms (words) had the same weights. Subsequently, Burrows described the “modified collection frequency weighing technique” in which each term (word) is weighted depending on what fraction of all documents the term occurs in so that, for example, an article or preposition that occurs in almost all documents would have a low weight. The “modified collection frequency weighing technique” of Burrows is described in more detail in Col. 26, lines 34 – 42. It is apparent from Burrows that this “modified collection frequency weighing technique” of Burrows is chosen at the inception and construction of the Burrows search engine (i.e., is a design choice of the search engine). The Burrows weighting scheme of a word is not modified during the operation of the search engine and

is “modified” only in that the technique implemented in the search engine of Burrows is a modification from the scheme proposed earlier in Burrows. In addition, the term weights described in Burrows are not modified during the operation of the search engine and certainly are not modified based on captured user actions as claimed.

#### Arguments

##### **Claims 1-6**

Ballard does not anticipate independent Claims 1 and 4. In particular, Claims 1 and 4 recite, a system and method, respectively, for user behavior based ranking of a document. Each claim recites determining a feature vector associated with a document wherein the feature vector comprises weights for certain terms that appear in the document, and modifying the feature vector for the document based on user actions captured during a search session so that the document is more highly ranked in response to the user actions. Ballard does not modify a feature vector of a document as claimed. As set forth above, Ballard refines a query (by adding an initial query together with one or more keywords) which is entirely different from the claimed modification of a feature vector of a document. In particular, the refinement of the query in Ballard permits the individual user to achieve a better search with the refined query whereas the claimed modification of the feature vector of the document results in the document having a different ranking for any subsequent user who might retrieve that document during a search. The claimed modification of the document feature vector permanently changes the feature vector of the document for all subsequent users of the system and is not simply refining the query of a particular user (like Ballard) which would have no effect on subsequent users. Therefore, this element of the claims is not present in Ballard.

Furthermore, Ballard does not describe modifying the feature vector for the document based on user actions captured during a search session as claimed. Ballard teaches that the user must explicitly enter classifications at a user interface and then keywords from the “of interest” documents are used to refine a query. In contrast, the claimed system and method captures user actions during a search session (such as the viewing of documents while browsing search results as described in the specification) in order to modify the feature vectors of the document. Thus, Ballard requires explicit user actions to classify documents (and at least one document must be classified as “of interest” and at least one document must be classified as “not of interest”) and then generates keywords while the claimed system and method captures user actions (and

therefore does not require the explicit classification actions by the user.) Therefore, this element of the claims is not described by Ballard. Therefore, Claims 1 and 4 are not anticipated by Ballard and are allowable over the prior art.

In addition, dependent Claims 2-3 and 5-6 (which depend from Claims 1 and 4) are patentable over Ballard for at least the same reasons as the independent claims. Furthermore, Claims 2 and 5 recite additional features not described in Ballard. In particular, Claim 2 recites that the capturing of the user actions includes selecting a document from the list of documents. Ballard describes a system in which a user must explicitly classify document since Ballard requires at least one document to be classified as "of interest" and at least one document classified as "not of interest" and either does not consider unclassified documents or classifies them as "not of interest." Therefore, Ballard does not describe this feature of the claimed invention and Claims 2 and 5 are allowable.

Claims 3 and 6 are not anticipated by Ballard. In particular, Ballard does not describe the feature recited in Claim 3 and 6. In particular, Ballard does not describe the claimed adjusting the weights of the terms in the feature vector that match terms in a query that produced the list of documents so that the ranking of the document is higher in response to the adjustment of the weights. Therefore, Claims 3 and 6 are allowable.

#### **Claims 7 - 12**

Ballard and Burrows, in combination, do not render Claims 7- 12 obvious under 35 USC 103. In particular, with respect to Claims 7 and 10 (the independent system and method claims), Ballard and Burrows in combination do not describe the features recited in these claims. For example, each of these claims recite "determining a feature vector associated with a document, the feature vector comprising weights for certain terms that appear in the document" and "modifying the feature vector for the document based on user actions captured during a query of the document so that the document is more highly ranked in response to the user actions." Ballard does not modify a feature vector of a document as claimed. As set forth above, Ballard refines a query (by adding an initial query together with one or more keywords) which is entirely different from the claimed modification of a feature vector of a document. In particular, the refinement of the query in Ballard permits the individual user to achieve a better search with the refined query whereas the claimed modification of the feature vector of the document results in the document having a different ranking for any subsequent user who might retrieve that

document during a search. The claimed modification of the document feature vector permanently changes the feature vector of the document for all subsequent users of the system and is not simply refining the query of a particular user (like Ballard) which would have no effect on subsequent users. Therefore, this element of the claims is not present in Ballard.

Furthermore, Ballard does not describe modifying the feature vector for the document based on user actions captured during a search session as claimed. Ballard teaches that the user must explicitly enter classifications at a user interface and then keywords from the “of interest” documents are used to refine a query. In contrast, the claimed system and method captures user actions during a search session (such as the viewing of documents while browsing search results as described in the specification) in order to modify the feature vectors of the document. Thus, Ballard requires explicit user actions to classify documents (and at least one document must be classified as “of interest” and at least one document must be classified as “not of interest”) and then generates keywords while the claimed system and method captures user actions (and therefore does not require the explicit classification actions by the user.) Therefore, these element of the claims are not described by Ballard.

Furthermore, Claims 7 and 10 recite “returning the same document to another user with the same query at a higher ranking due to the modified feature vector.” Ballard does not describe this feature of the claims as admitted by the Examiner. The Examiner has asserted that Burrows describes this feature of the claims. Burrows describes a technique in which a term is weighted depending on the fraction of all documents in which that the term occurs (which Burrows refers to as the “modified frequency collection weighing.”) For example, an article (“the”) appears in almost all documents and would be given a low term weight in Burrows for all queries and documents. The term weight assigned in Burrows does not change during the operation of the search engine and is not changed in response to user actions as claimed. Therefore, Burrows does not describe a claimed modified feature vector wherein the feature vector of the document is modified based on captured user actions over time, such as the viewing of documents. Because Burrows does not describe the claimed modified feature vector, Burrows also does not describe a system in which the claimed modified feature vector affects the search results of another user by ranking the document with the modified feature vector in a different manner. Therefore, Claims 7 and 10 are allowable over Ballard and Burrows.

With respect to Claims 8 and 11, these claims are allowable over Ballard and Burrows for at least the same reasons as the independent claims (Claims 7 and 10) as well as for at least the same reasons as Claim 2 above.

With respect to Claims 9 and 12, these claims are allowable over Ballard and Burrows for at least the same reasons as Claims 7 and 10 above as well as for at least the same reasons as Claims 3 and 6 above.

#### **New Claims 13 - 26**

New independent Claims 13 and 20 recite a method and system, respectively, for user behavior based ranking of a document. Ballard and Burrows, alone or in combination, do not anticipate these claims and therefore these claims are allowable for the reasons set forth below. In particular, the claims recite, in relevant part, sampling user search behavior and updating the feature vector of the document based on the sampled user search behavior so that the rank of the document is changed based on the user sampled user search behavior. These features of the claimed invention are not described in Ballard or Burrows. Ballard or Burrows does not describe sampling user action as claimed. Ballard teaches that the user must explicitly classify at least one "of interest" document and at least one "not of interest" document and that non categorized documents are ignored. Burrows does not sample user actions. Ballard or Burrows also does not describe updating the feature vector of the document. As set forth above, Ballard teaches modifying the query and not the claimed modification of the feature vector of the document. Burrows also does not describe the claimed feature vector modification. Therefore, Claims 13 and 20 are allowable.

Claims 14 and 21 are not anticipated by Ballard (or rendered obvious by Ballard and Burrows) for at least the same reasons as Claims 13 and 20 above. Furthermore, these claims recite generating a sample of the user behavior wherein the sample of the user behavior further comprises a query feature vector of the terms in a particular query and the feature vector of the one or more documents returned based on the query and viewed by the user. Ballard does not teach that the sample of the user behavior is the query feature vector and the feature vector of the one or more documents returned based on the query as claimed. At most, Ballard teaches the explicit user actions of classifying the documents. Therefore, Claims 14 and 21 are allowable.

Claims 15 and 22 are not anticipated by Ballard (or rendered obvious by Ballard and Burrows) for at least the same reasons as Claims 13 and 20 above. Furthermore, these claims

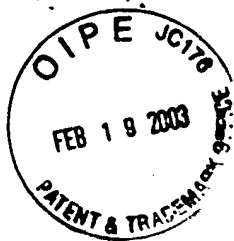
recite that the sample generating further comprises generating a sample during a sampling frequency. Ballard does not describe this feature of the invention. Therefore Claims 15 and 22 are allowable.

Claims 16 and 23 are not anticipated by Ballard (or rendered obvious by Ballard and Burrows) for at least the same reasons as Claims 13 and 20 above. Furthermore, these claims recite that the updating further comprises combining the feature vector of the document with a feature vector of the query, the feature vector comprising frequency values for one or more terms that appear in the query. Ballard does not describe this feature. Therefore, Claims 16 and 23 are allowable.

Claims 17 and 24 are not anticipated by Ballard (or rendered obvious by Ballard and Burrows) for at least the same reasons as Claims 16 and 23 above. Furthermore, these claims recite that the updating further comprising scaling the query feature vector based on the viewing time of the document by the user during the sampled user behavior to generate a scaled query feature vector. Ballard does not describe this feature. Therefore, Claims 17 and 24 are allowable.

Claims 18 and 25 are not anticipated by Ballard (or rendered obvious by Ballard and Burrows) for at least the same reasons as Claims 17 and 24 above. Furthermore, these claims recite that the scaling further comprises generating a negative scaling factor in response to a short viewing time so that the scaled query feature vector is negative and the feature vector of the document is reduced and the rank of the document is reduced. Ballard does not teach this feature. Therefore, Claims 18 and 25 are allowable.

Claims 19 and 26 are not anticipated by Ballard (or rendered obvious by Ballard and Burrows) for at least the same reasons as Claims 17 and 24 above. Furthermore, these claims recite that the scaling further comprises generating a positive scaling factor in response to a long viewing time so that the scaled query feature vector is positive and the feature vector of the document is increased and the rank of the document is increased. Ballard does not teach this feature. Therefore, Claims 19 and 26 are allowable.



CONCLUSION

In view of the above, it is respectfully submitted that Claims 1- 26 are allowable over the prior art cited by the Examiner and early allowance of these claims and the application is respectfully requested.

The Examiner is invited to call Applicant's attorney at the number below in order to speed the prosecution of this application.

The Commissioner is authorized to charge any deficiencies in fees and credit any overpayment of fees to Deposit Account No. 07-1896.

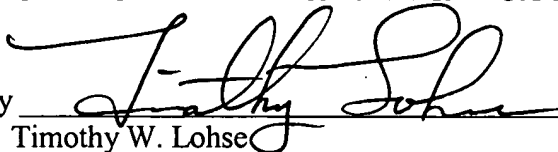
Respectfully submitted,

GRAY CARY WARE & FREIDENRICH LLP

Dated: \_\_\_\_\_

February 11, 2003

By \_\_\_\_\_



Timothy W. Lohse  
Reg. No. 35,255  
Attorney for Applicant

GRAY CARY WARE & FREIDENRICH  
1755 Embarcadero Rd,  
Palo Alto, CA 94303-3340  
Telephone: (650) 320-7426





APPENDIX A  
MARKED UP CLAIMS

Claims:

- 1.(Amended) A system for user behavior based ranking of a document, comprising:  
means for determining a feature vector associated with a document, the feature vector comprising weights for certain terms that appear in the document; and  
means for modifying the feature vector for the document based on user actions captured during a search session so that the document is more highly ranked in response to the user actions.
- 2.(Amended) The system of Claim 1 further comprising means for [collecting] capturing user actions in response to a list of documents produced in response to a query wherein the user actions include selecting a document from the list of documents.
- 4.(Amended) A method for user behavior based ranking of a document, comprising:  
determining a feature vector associated with a document, the feature vector comprising weights for one or more terms that appear in the document; and  
modifying the feature vector for the document based on user actions captured during a query of the document so that the document is more highly ranked in response to the user actions.
5. (Amended) The method of Claim 4 further comprising means for [collecting] capturing user actions in response to a list of documents produced in response to a query wherein the user actions include selecting a document from the list of documents.
- 7.(Amended) A system for user behavior based searching of a document based on a query having one or more query terms, comprising:

RECEIVED  
FEB 24 2003  
Technology Center 2100

means for determining a feature vector associated with a document, the feature vector comprising weights for certain terms that appear in the document;

means for modifying the feature vector for the document based on user actions captured during a query of the document so that the document is more highly ranked in response to the user actions; and

means for returning the same document to another user with the same query at a higher ranking due to the modified feature vector.

8.(Amended) The system of Claim 7 further comprising means for [collecting] capturing user actions in response to a list of documents produced in response to a query wherein the user actions include selecting a document from the list of documents.

10. (Amended) A method for user behavior based searching of a document based on a query having one or more query terms, comprising:

determining a feature vector associated with a document, the feature vector comprising frequency values for one or more terms that appear in the document;

modifying the feature vector for the document based on user actions captured during a query of the document so that the document is more highly ranked in response to the user actions; and

returning the same document to another user with the same query at a higher ranking due to the modified feature vector.

11. (Amended) The method of Claim 10 further comprising means for [collecting] capturing user actions in response to a list of documents produced in response to a query wherein the user actions include selecting a document from the list of documents.